

CLAIMS

1. A method of detecting of high-grade dysplasia (HGD) in cells of a tissue sample, the method comprising:

- (a) obtaining a test tissue sample suspected of comprising cells exhibiting HGD;
- (b) establishing the level of expression in the test tissue sample of at least eight genes selected from the group consisting of ET-1 (endothelin-1, NM_001955) (SEQ ID NO:1); AGR2 (anterior gradient 2 (*Xenopus laevis*) homolog, NM_006408) (SEQ ID NO:3); ADAM8 (NM_001109) (SEQ ID NO:5); PRSS8 (Prostasin precursor, serine protease, NM_002773) (SEQ ID NO:7); AXO1 (Axonin-1 precursor, NM_005076) (SEQ ID NO:9); NROB2 (Nuclear hormone receptor, NM_021969) (SEQ ID NO:11); TM7SF1 (NM_003272) (SEQ ID NO:13); DLDH (dihydrolipamide dehydrogenase, NM_000108) (SEQ ID NOS:15); MAT2B (methionine adenosyltransferase II, beta, NM_013283) (SEQ ID NO:17); STC-2 (stanniocalcin-2, NM_003714) (SEQ ID NO:19); PPBI (alkaline phosphatase, intestinal precursor, NM_001631) (SEQ ID NO:21); SLNAC1 (sodium channel receptor SLNAC1, NM_004769) (SEQ ID NO:23); CAH4 (carbonic anhydrase iv precursor, NM_000717) (SEQ ID NO:25); PA21 (phospholipase a2 precursor, NM_000928) (SEQ ID NO:27); PAR2 (proteinase activated receptor 2 precursor, NM_005242) (SEQ ID NO:29); IDE (insulin-degrading enzyme, NM_004969) (SEQ ID NO:31); MYO1A (myosin-1A, NM_005379) (SEQ ID NO:33); CYP2J2 (cytochrome P450 monooxygenase, NM_000775) (SEQ ID NO:35); PHYH (phytanoyl-CoA-hydroxylase (Refsum disease), NM_006214) (SEQ ID NO:37); CYB5 (cytochrome b5, 3' end, NM_001914) (SEQ ID NO:39); COXVIb (coxVIb gene, last exon and flanking sequence, NM_001863) (SEQ ID NO:41); and TCF4 (NM_030756) (SEQ ID NO:43), or variants thereof having at least 80% nucleic acid sequence identity, wherein the tissue is from esophagus or colon; and
- (c) comparing expression of the at least eight genes to a baseline expression of the genes in normal tissue controls of the same tissue type, wherein an increase of at least 1.5-fold in expression of the genes relative to the baseline expression indicates that cells of the test sample exhibit HGD.

2. The method of claim 1, wherein the tissue is human tissue.

3. A method of identifying a esophageal tissue susceptible to esophageal adenocarcoma, comprising detecting esophageal HGD in a test tissue sample according to claim 1.
4. A method according to claim 1, wherein an increase of at least 2-fold in expression of genes relative to the baseline is observed.
5. A method according to claim 1, wherein at least one of the at least eight genes is selected from the group consisting of AGR2 (SEQ ID NO:3), TM7SF1 (SEQ ID NO:13), MAT2B (SEQ ID NO:17), SLNAC1 (SEQ ID NO:23), and TCF4 (SEQ ID NO:43), or variants thereof having at least 80% nucleic acid sequence identity.
6. A method for determining predisposition of a mammalian tissue to a neo-plastic transformation by detecting HGD in cells of the tissue, the method comprising determining in a cell from the tissue expression of a nucleic acid sequence of at least eight genes selected from the group consisting of ET-1 (endothelin-1, NM_001955) (SEQ ID NO:1); AGR2 (anterior gradient 2 (*Xenopus laevis*) homolog, NM_006408) (SEQ ID NO:3); ADAM8 (NM_001109) (SEQ ID NO:5); PRSS8 (Prostasin precursor, serine protease, NM_002773) (SEQ ID NO:7); AXO1 (Axonin-1 precursor, NM_005076) (SEQ ID NO:9); NROB2 (Nuclear hormone receptor, NM_021969) (SEQ ID NO:11); TM7SF1 (NM_003272) (SEQ ID NO:13); DLDH (dihydrolipamide dehydrogenase, NM_000108) (SEQ ID NOS:15); MAT2B (methionine adenosyltransferase II, beta, NM_013283) (SEQ ID NO:17); STC-2 (stanniocalcin-2, NM_003714) (SEQ ID NO:19); PPBI (alkaline phosphatase, intestinal precursor, NM_001631) (SEQ ID NO:21); SLNAC1 (sodium channel receptor SLNAC1, NM_004769) (SEQ ID NO:23); CAH4 (carbonic anhydrase iv precursor, NM_000717) (SEQ ID NO:25); PA21 (phopholipase a2 precursor, NM_000928) (SEQ ID NO:27); PAR2 (proteinase activated receptor 2 precursor, NM_005242) (SEQ ID NO:29); IDE (insulin-degrading enzyme, NM_004969) (SEQ ID NO:31); MYO1A (myosin-1A, NM_005379) (SEQ ID NO:33); CYP2J2 (cytochrome P450 monooxygenase, NM_000775) (SEQ ID NO:35); PHYH (phytanoyl-CoA-hydroxylase (Refsum disease), NM_006214) (SEQ ID NO:37); CYB5 (cytochrome b5, 3' end, NM_001914) (SEQ ID NO:39); COXVIb (coxVIb gene, last exon and flanking sequence, NM_001863) (SEQ ID NO:41); and TCF4 (NM_030756) (SEQ ID NO:43), or variants thereof having at least 80%

nucleic acid sequence identity, wherein the tissue of from esophagus or colon, and wherein the expression in the test sample is at least 1.5-fold above baseline expression in a normal tissue control of the same tissue type.

7. A method according to claim 6, wherein the tissue is human tissue.

8. A method according to claim 6, wherein at least one of the at least eight genes is selected from the group consisting of AGR2 (SEQ ID NO:3), TM7SF1 (SEQ ID NO:13), MAT2B (SEQ ID NO:17), SLNAC1 (SEQ ID NO:23), and TCF4 (SEQ ID NO:43), or variants thereof having at least 80% nucleic acid sequence identity.

9. A method of detecting high-grade dysplasia (HGD) in cells of a mammalian tissue sample, the method comprising:

- (a) obtaining a test tissue sample suspected of comprising cells exhibiting HGD;
- (b) establishing the level of expression in the test tissue sample of at least eight polypeptides encoded by genes selected from the group consisting of ET-1 (endothelin-1, NM_001955) (SEQ ID NO:1); AGR2 (anterior gradient 2 (*Xenopus laevis*) homolog, NM_006408) (SEQ ID NO:3); ADAM8 (NM_001109) (SEQ ID NO:5); PRSS8 (Prostasin precursor, serine protease, NM_002773) (SEQ ID NO:7); AXO1 (Axonin-1 precursor, NM_005076) (SEQ ID NO:9); NROB2 (Nuclear hormone receptor, NM_021969) (SEQ ID NO:11); TM7SF1 (NM_003272) (SEQ ID NO:13); DLDH (dihydrolipamide dehydrogenase, NM_000108) (SEQ ID NOS:15); MAT2B (methionine adenosyltransferase II, beta, NM_013283) (SEQ ID NO:17); STC-2 (stanniocalcin-2, NM_003714) (SEQ ID NO:19); PPBI (alkaline phosphatase, intestinal precursor, NM_001631) (SEQ ID NO:21); SLNAC1 (sodium channel receptor SLNAC1, NM_004769) (SEQ ID NO:23); CAH4 (carbonic anhydrase iv precursor, NM_000717) (SEQ ID NO:25); PA21 (phopholipase a2 precursor, NM_000928) (SEQ ID NO:27); PAR2 (proteinase activated receptor 2 precursor, NM_005242) (SEQ ID NO:29); IDE (insulin-degrading enzyme, NM_004969) (SEQ ID NO:31); MYO1A (myosin-1A, NM_005379) (SEQ ID NO:33); CYP2J2 (cytochrome P450 monooxygenase, NM_000775) (SEQ ID NO:35); PHYH (phytanoyl-CoA-hydroxylase (Refsum disease), NM_006214) (SEQ ID

NO:37); CYB5 (cytochrome b5, 3' end, NM_001914) (SEQ ID NO:39); COXVIb (coxVIb gene, last exon and flanking sequence, NM_001863) (SEQ ID NO:41); and TCF4 (NM_030756) (SEQ ID NO:43), or variants thereof having at least 80% nucleic acid sequence identity, wherein the tissue is from esophagus or colon; and

(c) comparing expression of the at least eight polypeptides in the test tissue sample to expression of the at least eight polypeptides in normal tissue controls of the same tissue type, wherein an increase of at least 1.5-fold in expression of the polypeptides in the test tissue sample relative to the normal tissue controls indicates that cells of the test sample exhibit HGD.

10. A method as according to claim 9 comprising contacting the test tissue sample with an antibody that specifically binds one of the at least eight polypeptides under conditions that permit the antibody to bind the polypeptide.

11. A method according to claim 9, wherein at least one of the at least eight polypeptides expressed by a gene selected from the group consisting of AGR2 (SEQ ID NO:3), TM7SF1 (SEQ ID NO:13), MAT2B (SEQ ID NO:17), SLNAC1 (SEQ ID NO:23), and TCF4 (SEQ ID NO:43), or variants thereof having at least 80% nucleic acid sequence identity.

12. The method of claim 1, wherein gene expression is determined by nucleic acid microarray analysis.

13. The method of claim 12, wherein analysis comprises contacting nucleic acid from a test tissue sample with a nucleic acid microarray comprising nucleic acid probe sequences, wherein at least eight of the nucleic acid probe sequences separately comprises at least 50 contiguous nucleotides from a gene selected from the group consisting of ET-1 (endothelin-1, NM_001955) (SEQ ID NO:1); AGR2 (anterior gradient 2 (*Xenopus laevis*) homolog, NM_006408) (SEQ ID NO:3); ADAM8 (NM_001109) (SEQ ID NO:5); PRSS8 (Prostasin precursor, serine protease, NM_002773) (SEQ ID NO:7); AXO1 (Axonin-1 precursor, NM_005076) (SEQ ID NO:9); NROB2 (Nuclear hormone receptor, NM_021969) (SEQ ID NO:11); TM7SF1 (NM_003272) (SEQ ID NO:13); DLDH (dihydrolipamide dehydrogenase, NM_000108) (SEQ ID NOS:15);

MAT2B (methionine adenosyltransferase II, beta, NM_013283) (SEQ ID NO:17); STC-2 (stanniocalcin-2, NM_003714) (SEQ ID NO:19); PPBI (alkaline phosphatase, intestinal precursor, NM_001631) (SEQ ID NO:21); SLNAC1 (sodium channel receptor SLNAC1, NM_004769) (SEQ ID NO:23); CAH4 (carbonic anhydrase iv precursor, NM_000717) (SEQ ID NO:25); PA21 (phospholipase a2 precursor, NM_000928) (SEQ ID NO:27); PAR2 (proteinase activated receptor 2 precursor, NM_005242) (SEQ ID NO:29); IDE (insulin-degrading enzyme, NM_004969) (SEQ ID NO:31); MYO1A (myosin-1A, NM_005379) (SEQ ID NO:33); CYP2J2 (cytochrome P450 monooxygenase, NM_000775) (SEQ ID NO:35); PHYH (phytanoyl-CoA-hydroxylase (Refsum disease), NM_006214) (SEQ ID NO:37); CYB5 (cytochrome b5, 3' end, NM_001914) (SEQ ID NO:39); COXVIb (coxVIb gene, last exon and flanking sequence, NM_001863) (SEQ ID NO:41); and TCF4 (NM_030756) (SEQ ID NO:43), or variants thereof having at least 80% nucleic acid sequence identity..

14. The method of claim 13, wherein the at least eight nucleic acid probe sequences comprise at least 60 contiguous nucleotides from a gene selected from the group.

15. The method of claim 14, wherein the at least eight nucleic acid probe sequences comprise at least 80 contiguous nucleotides from a gene selected from the group.

16. The method of claim 15, wherein the at least eight nucleic acid probe sequences comprise at least 100 contiguous nucleotides from a gene selected from the group.

17. The method of claim 16, wherein the at least eight nucleic acid probe sequences comprise at least 150 contiguous nucleotides from a gene selected from the group.

18. The method of claim 17, wherein the at least eight nucleic acid probe sequences comprise at least 200 contiguous nucleotides from a gene selected from the group.

19. The method of claim 13, wherein the nucleic acid microarray comprises nucleic acid probe sequences from at least ten genes selected from the group.

20. The method of claim 19, wherein the nucleic acid microarray comprises nucleic acid probe sequences from at least twelve genes selected from the group.
21. The method of claim 20, wherein the nucleic acid microarray comprises nucleic acid probe sequences from at least fifteen genes selected from the group.
22. The method of claim 21, wherein the nucleic acid microarray comprises nucleic acid probe sequences from at least eighteen genes selected from the group.
23. The method of claim 22, wherein the nucleic acid microarray comprises nucleic acid probe sequences from at least twenty genes selected from the group.
24. The method of claim 23, wherein the nucleic acid microarray comprises nucleic acid probe sequences from at least twenty two genes selected from the group.
25. The method of claim 1, wherein gene expression is determined by nucleic acid hybridization under high stringency conditions of a detectable probe comprising at least 50 contiguous nucleotides from a gene selected from the group to nucleic acid of cells of the test tissue sample relative to cells of the normal tissue control.
26. The method of claim 25, wherein the hybridization is *in situ* hybridization.
27. The method of claim 26, wherein the hybridization is fluorescent *in situ* hybridization.
28. The method of claim 1, wherein gene expression is determined by polymerase chain reaction (PCR) analysis.
29. The method of claim 1, wherein gene expression is determined by real-time polymerase chain reaction (RT-PCR) analysis.

30. The method of claim 1, wherein gene expression is determined by Taqman® polymerase chain reaction analysis.

31. A kit comprising a microarray, the microarray comprising nucleic acid probe sequences, wherein at least eight of the nucleic acid probe sequences each comprise at least 50 contiguous nucleotides from a gene selected from the group consisting of ET-1 (endothelin-1, NM_001955) (SEQ ID NO:1); AGR2 (anterior gradient 2 (*Xenopus laevis*) homolog, NM_006408) (SEQ ID NO:3); ADAM8 (NM_001109) (SEQ ID NO:5); PRSS8 (Prostasin precursor, serine protease, NM_002773) (SEQ ID NO:7); AXO1 (Axonin-1 precursor, NM_005076) (SEQ ID NO:9); NROB2 (Nuclear hormone receptor, NM_021969) (SEQ ID NO:11); TM7SF1 (NM_003272) (SEQ ID NO:13); DLDH (dihydrolipamide dehydrogenase, NM_000108) (SEQ ID NOS:15); MAT2B (methionine adenosyltransferase II, beta, NM_013283) (SEQ ID NO:17); STC-2 (stanniocalcin-2, NM_003714) (SEQ ID NO:19); PPBI (alkaline phosphatase, intestinal precursor, NM_001631) (SEQ ID NO:21); SLNAC1 (sodium channel receptor SLNAC1, NM_004769) (SEQ ID NO:23); CAH4 (carbonic anhydrase iv precursor, NM_000717) (SEQ ID NO:25); PA21 (phospholipase a2 precursor, NM_000928) (SEQ ID NO:27); PAR2 (proteinase activated receptor 2 precursor, NM_005242) (SEQ ID NO:29); IDE (insulin-degrading enzyme, NM_004969) (SEQ ID NO:31); MYO1A (myosin-1A, NM_005379) (SEQ ID NO:33); CYP2J2 (cytochrome P450 monooxygenase, NM_000775) (SEQ ID NO:35); PHYH (phytanoyl-CoA-hydroxylase (Refsum disease), NM_006214) (SEQ ID NO:37); CYB5 (cytochrome b5, 3' end, NM_001914) (SEQ ID NO:39); COXVIb (coxVIb gene, last exon and flanking sequence, NM_001863) (SEQ ID NO:41); and TCF4 (NM_030756) (SEQ ID NO:43), or variants thereof having at least 80% nucleic acid sequence identity, and a package insert indicating that the microarray is for use in detecting HGD in a test tissue sample, wherein the tissue is from esophagus or colon, and wherein an increase in expression in the test tissue sample of at least 1.5-fold of the at least eight genes relative to a normal tissue control of the same tissue type indicates that cells of the test tissue exhibit HGD.

32. The kit of claim 31, wherein the nucleic acid probe sequences each comprise at least 60 contiguous nucleotides from a gene selected from the group.

33. The kit of claim 32, wherein the nucleic acid probe sequences each comprise at least 80 contiguous nucleotides from a gene selected from the group.
34. The kit of claim 33, wherein the nucleic acid probe sequences each comprise at least 100 contiguous nucleotides from a gene selected from the group.
35. The kit of claim 34, wherein the nucleic acid probe sequences each comprise at least 150 contiguous nucleotides from a gene selected from the group.
36. The kit of claim 35, wherein the nucleic acid probe sequences each comprise at least 200 contiguous nucleotides from a gene selected from the group.
37. A method of detecting cancer in a patient, the method comprising:
- (a) obtaining a test tissue sample from the patient;
 - (b) establishing the level of expression of a gene selected from the group consisting of CAD17 (liver-intestine cadherin, NM_004063) (SEQ ID NO:45), CLDN15 (claudin 15, NM_014343) (SEQ ID NO:47), SLNAC1 (sodium channel, NM_004769) (SEQ ID NO:23), CFTR (chloride channel, NM_000492) (SEQ ID NO:49), H2R (histamine H2 receptor, NM_022304) (SEQ ID NO:51), PRSS8 (serine protease, NM_002773) (SEQ ID NO:7), PA21 (phospholipase A2 group IB, NM_000928) (SEQ ID NO:27), AGR2 (anterior gradient 2 homolog, (NM_006408) (SEQ ID NO:3), EGFR (NM_005228) (SEQ ID NO:53), EPHB2 (NM_004442) (SEQ ID NO:55), CRIPTO CR-1 (NM_003212) (SEQ ID NO:57), Eprin B1 (NM_004429) (SEQ ID NO:59), MMP-17/MT4-MMP (NM_016155) (SEQ ID NO:61), MMP26 (NM_021801) (SEQ ID NO:63), ADAM10 (NM_001110) (SEQ ID NO:65), ADAM8 (NM_001109) (SEQ ID NO:5), ADAM1 (XM_132370) (SEQ ID NO:67), TIM1 (NM_003254) (SEQ ID NO:69), MUC1 (XM_053256) (SEQ ID NO:71), CEA (NM_004363) (SEQ ID NO:73), NCA (NM_002483) (SEQ ID NO:75), Follistatin (NM_006350) (SEQ ID NO:77), Claudin 1 (NM_021101) (SEQ ID NO:79), Claudin 14 (NM_012130) (SEQ ID NO:81), tenascin-R (NM_003285) (SEQ ID NO:83), CAD3 (NM_001793) (SEQ ID NO:85), AXO1 (NM_005076) (SEQ ID NO:9), CONT (NM_001843) (SEQ ID NO:87), Osteopontin (NM_000582) (SEQ ID NO:89), Galectin 8 (NM_006499) (SEQ ID NO:91), PGS1 (bihlycan,

NM_001711) (SEQ ID NO:93), Frizzled 2 (NM_001466) (SEQ ID NO.:95), ISLR (NM_005545) (SEQ ID NO:97), FLJ23399 (NM_022763) (SEQ ID NO:99), TEM1 (NM_020404) (SEQ ID NO:101), Tie2 ligand2 (NM_001147) (SEQ ID NO:103), STC-2 (NM_003714) (SEQ ID NO:19), VEGFC (NM_005429) (SEQ ID NO:105), tPA (NM_000930) (SEQ ID NO:107), Endothelin 1 (NM_001955) (SEQ ID NO:1), Thrombomodulin (NM_000361) (SEQ ID NO:109), TF (NM_001993) (SEQ ID NO:111), GPR4 (NM_005282) (SEQ ID NO:113), GPR66 (NM_006056) (SEQ ID NO:115), SLC22A2 (NM_003058) ((SEQ ID NO:117), MLSN1 (NM_002420) (SEQ ID NO:119), and ATN2 (Na/K transport, NM_000702) (SEQ ID NO:121), or variants thereof having at least 80% nucleic acid sequence identity, wherein the test tissue is from esophagus or colon; and wherein the expressing in the test tissue is at a level at least 1.5-fold above expression of the gene in a normal tissue control of the same tissue type.

38. The method of claim 37, wherein inhibition of cell growth is cell death.

39. The method of claim 37, wherein at least two genes selected from the group are expressed at a level at least 1.5-fold above expression of the gene in a normal cell control.

40. The method of claim 39, wherein at least three genes selected from the group are expressed at a level at least 1.5-fold above expression of the gene in a normal cell control.

41. The method of claim 40, wherein at least 5 genes selected from the group are expressed at a level at least 1.5-fold above expression of the gene in a normal cell control.

42. The method of claim 41, wherein at least 8 genes selected from the group are expressed at a level at least 1.5-fold above expression of the gene in a normal cell control.

43. The method of claim 1, wherein the expression p value is less than 0.07.

44. The method of claim 6, wherein the expression p value is less than 0.07.

P2000R1

45. The method of claim 9, wherein the expression p value is less than 0.07.